

G-coupling functions: The Infinite Dimensional case

D.M. Morales Silva

School of ITMS, University of Ballarat, VIC 3353, Australia
dmoralessilva@students.ballarat.edu.au

Abstract

In this work we present a class of functions, motivated by gap functions, which we call G-coupling functions. We will show that these functions can generate a duality scheme for minimization problems by means of the general conjugation theory. Thanks to this scheme, a Lagrange-type function is introduced as well.

Keywords: general conjugation theory, non convex optimization, gap functions.

1 Introduction

For solving non-convex optimization problems, a tool that is becoming more important is *generalized conjugation*. In [4] the G-coupling functions are introduced in finite dimensional spaces. Here we extend this definition to the infinite dimensional case. These coupling functions will allow us to see duality schemes in a different way. The usual theory found in the literature ([6], [8], and references therein) are related to a fixed coupling function, but here we consider (for a specified function f) a family of coupling functions.

These coupling functions are motivated by gap functions. It is interesting to point out, that many of these (gap) functions have similar properties. However, in some cases they are functions of one vector and it is important, since they are linked to specified optimization problems, that those functions have zeros.

On the other hand, G-coupling functions will be defined as functions in two variables and they might not have zeros. Even more, given a specified proper function f , it is shown that a certain sub-family of this family of coupling functions satisfies many interesting properties.

In Section 2, we describe how many gap functions have similar properties, which are useful for the definition of G-coupling functions.

In Section 3, it is found the definition of G-coupling function with properties related to generalized conjugation using this family of functions and a fixed proper function f .

In Section 4, it can be seen how these ideas generate Lagrange-type functions (see [7]).

2 Motivation

In several works already published, there can be found definitions of GAP functions for particular problems. Now we present 2 concrete examples.

In [2], the Variational Inequality Problem is studied:

$$(VIP) \text{ Find } x_0 \in C, \text{ such that, } \exists y^* \in T(x_0) \text{ with } \langle y^*, x - x_0 \rangle \geq 0 \quad \forall x \in C,$$

where T is a maximal monotone correspondence which is defined as follows: given a point to set map, T , it will be said that it is a maximal monotone correspondence if it satisfies that $\langle u - v, x - y \rangle \geq 0$ for every $u \in T(x)$, $v \in T(y)$ with $x, y \in C$ and if there exists v ,

such that $\langle u - v, x - y \rangle \geq 0$, for all $x, y \in C$ and for all $u \in T(x)$, then $v \in T(y)$. The corresponding GAP function is then defined as follows:

$$h_{T,C}(x) := \sup_{(v,y) \in G_C(T)} \langle v, x - y \rangle,$$

where $G_C(T) = \{(v, y) : v \in T(y), y \in C\}$ and C is a non-empty closed convex set. This function happens to be non-negative and convex, and it is equal to zero only in solutions of (VIP).

In [8], the Equilibrium Problem is studied:

$$(EP) \text{ Find } x \in K, \text{ such that } f(x, y) \geq 0, \forall y \in K,$$

where $K \subset \mathbb{R}^n$ is a non-empty closed convex set and $f : K \times K \rightarrow \mathbb{R}$ is a function that satisfies:

- i) $f(x, x) = 0$, for all $x \in K$.
- ii) $f(x, \cdot) : K \rightarrow \mathbb{R}$ is convex and l.s.c.
- iii) $f(\cdot, y) : K \rightarrow \mathbb{R}$ is u.s.c.

The GAP function is defined as:

$$g_f(y) := \begin{cases} \sup_{x \in K} f(x, y) & \text{if } y \in K \\ +\infty & \text{in other case.} \end{cases}$$

In this case, the function g_f is non-negative, convex and l.s.c. and if it vanishes at x_0 , then x_0 is a solution of (EP).

In these examples, gap functions are used to transform a special Equilibrium Problem (for example, the VIP is a particular case of an EP) into a minimization problem.

Now our attention is focused in using coupling functions that could be related, at least in some general aspect, to GAP functions. Therefore these functions must link both primal and dual variables. Since these coupling functions must be related to a sense of “gap”, we consider these functions as non-negative and with 2 arguments.

Let us remember that for the minimization problem, the convex conjugation theory allows us to generate a dual problem and there is implicit another concept of gap function (see [1], [3] and [5]): consider

$$\alpha = \inf[f(x) : x \in \mathbb{R}^n]. \quad (P)$$

Define a function $\varphi : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$, satisfying

$$\varphi(x, 0) = f(x), \forall x \in \mathbb{R}^n.$$

Then φ will be called a perturbation function and the function $h : \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$ defined by

$$h(u) = \inf_{x \in \mathbb{R}^n} \varphi(x, u)$$

will be called the marginal function. Observe that

$$\alpha = h(0) = \inf_{x \in \mathbb{R}^n} \varphi(x, 0) = \inf_{x \in \mathbb{R}^n} f(x).$$

Considering now h^{**} , the convex bi-conjugate (see [5]) of h one has:

$$h^{**}(0) \leq h(0) = \alpha$$

where

$$h^{**}(0) = \sup[\langle u^*, 0 \rangle - h^*(u^*) : u^* \in \mathbb{R}^p].$$

Then, making $-\beta = h^{**}(0)$, one has

$$\beta = \inf_{u^* \in \mathbb{R}^p} h^*(u^*). \quad (Q)$$

(Q) is called the dual problem of (P) and in general we have $-\beta \leq \alpha$. It is said that there is no duality gap whenever $h^{**}(0) = h(0)$. It is easy to prove that $h^*(u^*) = \varphi^*(0, u^*)$, and if we define the function $k : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ by $k(x^*) := \inf_{u^* \in \mathbb{R}^p} \varphi^*(x^*, u^*)$, then $\beta = k(0)$.

This analysis is summarized in the following scheme:

$$\begin{array}{llll} \alpha & = & \inf f(x) & (P) \\ \varphi(x, 0) & = & f(x), \forall x \in \mathbb{R}^n & \\ h(u) & = & \inf_x \varphi(x, u) & \\ \alpha & = & h(0) & \end{array} \quad \begin{array}{llll} \beta & = & \inf h^*(u^*) & (Q) \\ \varphi^*(0, u^*) & = & h^*(u^*), \forall u^* \in \mathbb{R}^p & \\ k(x^*) & = & \inf_{u^*} \varphi^*(x^*, u^*) & \\ \beta & = & k(0) & \end{array}$$

$$-\beta \leq \alpha.$$

If h is proper and convex, a necessary and sufficient condition for ensuring that there will be no duality gap ($-\beta = \alpha$) is that h be l.s.c. at 0 (in general φ l.s.c. does not imply that h would be l.s.c.).

Further more, if h is convex, l.s.c. and $0 \in \text{ri}(\text{dom}(h))$, then $\alpha = -\beta$ and the dual problem has at least one optimal solution, and if $\overline{u^*}$ is an optimal solution of (Q) and $\varphi = \varphi^{**}$, then

$$\overline{x} \text{ is an optimal solution of } (P) \iff f(\overline{x}) + h^*(\overline{u^*}) = 0.$$

Consider now the function $g : \mathbb{R}^n \times \mathbb{R}^p \rightarrow \overline{\mathbb{R}}$ defined by:

$$g(x, u^*) = f(x) + h^*(u^*).$$

This function vanishes at (x_0, u_0^*) if and only if x_0 solves the primal problem and u_0^* solves the dual one. In addition, this function is non-negative and if the first variable is kept fixed, the function is convex and l.s.c. It is clear now, which properties are satisfied for many gap functions.

3 G-coupling Functions

As stated before, G-coupling functions are first introduced in [4] for finite dimensional spaces. We are going to extend this notion for arbitrary Banach spaces.

Henceforth, we consider two arbitrary Banach spaces X and Y .

Definition 3.1 A non-negative function $g : A \times B \rightarrow \mathbb{R}$, with $A \times B \subset X \times Y$ will be called a G-coupling function if

$$(D1) \quad \inf_{x \in A, y \in B} g(x, y) = 0.$$

Define

$$\mathcal{F}^{A,B} := \{g : A \times B \rightarrow \mathbb{R} : g \text{ is a G-coupling function}\}. \quad (1)$$

Not every G-coupling function has zeros:

Example: Define on $X \times Y$

$$g(x, y) = \exp(\|x\| - \|y\|).$$

Then $g \in \mathcal{F}^{X,Y}$ is continuous and it does not have any zeros.

Let us turn our attention now to how the family of functions $\mathcal{F}^{A,B}$ will allow us to establish duality schemes in (at least for now) the minimization problem. It is important to point out that in the following we consider an unusual type of duality, $f : A \rightarrow \mathbb{R} \cup \{+\infty\}$ is kept fixed and $g \in \mathcal{F}^{A,B}$, for a given $B \subset Y$, is variable.

Consider a proper function $f : A \rightarrow \mathbb{R} \cup \{+\infty\}$. For a given $B \subset Y$ take $g \in \mathcal{F}^{A,B}$. Define $f^g : B \rightarrow \mathbb{R} \cup \{+\infty\}$ and $f^{gg} : A \rightarrow \mathbb{R} \cup \{+\infty\}$ as follows (for example see [6] and references therein):

$$f^g(y) := \sup_{x \in A} \{g(x, y) - f(x)\} \quad \forall y \in B, \quad (2)$$

$$f^{gg}(x) := \sup_{y \in B} \{g(x, y) - f^g(y)\} \quad \forall x \in A. \quad (3)$$

In some cases, it would be better to consider a $g \in \mathcal{F}^{A,B}$ which satisfies:

(D2) B is convex and $g(x, \cdot) : B \rightarrow \mathbb{R}$ is a convex and l.s.c. function for each x in A .

With this, we have the following:

Lemma 3.1 *Let $f : A \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper function and given $B \subset Y$ take $g \in \mathcal{F}^{A,B}$. Then*

$$f^g(y) + f(x) \geq g(x, y) \geq 0, \quad \forall (x, y) \in A \times B, \quad (4)$$

which implies

$$f(x) \geq -f^g(y), \quad \forall (x, y) \in A \times B. \quad (5)$$

Moreover if g satisfies (D2), then f^g is a convex l.s.c function.

Unless it is mentioned, not every $g \in \mathcal{F}^{A,B}$ satisfies (D2).

It would be interesting to know which condition either a G-coupling function g or the function f must satisfy in order that the function f^g be proper, because with this one would have a non-trivial function related to f . The following lemma ensures the existence of such a function $g \in \mathcal{F}^{A,B}$ for any $B \subset Y$, taking as a starting point a natural condition on f which must be imposed if f is the objective function of a minimization problem.

Lemma 3.2 *Let f be as before. Then f is bounded from below if and only if, for every non-empty $B \subset Y$, there exists $g \in \mathcal{F}^{A,B}$ such that f^g is proper.*

Proof:

- Suppose that $\inf f > -\infty$, then for a non-empty $B_0 \subset Y$ fixed, consider $g \in \mathcal{F}^{A,B_0}$ as follows:

$$g(x, y) = \|y\|, \quad \forall (x, y) \in A \times B_0.$$

Thus

$$f^g(y) = \|y\| - \inf f \quad \forall y \in B_0,$$

which is clearly a proper function and since $B_0 \subset Y$ was fixed arbitrarily, the result is satisfied for every $B \subset Y$.

- Take a non-empty $B_0 \subset Y$ and $g \in \mathcal{F}^{A,B_0}$ such that f^g is proper. Let us suppose that $\inf f = -\infty$, from [6] we can see that this implies that $\inf f^{gg} = -\infty$. Then:

$$-\infty = \inf f^{gg} = \inf_{x \in A} \left(\sup_{y \in B_0} [g(x, y) - f^g(y)] \right) \geq$$

$$\sup_{y \in B_0} \left(\inf_{x \in A} [g(x, y) - f^g(y)] \right) \geq \sup_{y \in B_0} (-f^g(y)) = - \inf_{y \in B_0} f^g(y),$$

which means $-\infty \geq - \inf_{y \in B_0} (f^g(y))$. Then $\inf_{y \in B_0} f^g(y) = +\infty$, which implies that f^g is not proper and we have a contradiction. Therefore we must have that $\inf f > -\infty$.

Notice that this proof also states, in particular, that there exists $g \in \mathcal{F}^{A,B}$ for every non-empty $B \subset Y$ which satisfies (D2) and f^g is proper.

Given non-empty sets $A \subset X$ and $B \subset Y$, let

$$\mathcal{F}^A := \{f : A \rightarrow \mathbb{R} \cup \{+\infty\}, f \text{ is proper, } \inf f > -\infty\} \quad (6)$$

and $\gamma_{g,f} : A \times B \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by:

$$\gamma_{g,f}(x, y) := f(x) + f^g(y) \quad (7)$$

with $g \in \mathcal{F}^{A,B}$ and $f \in \mathcal{F}^A$. Take $f \in \mathcal{F}^A$ and define

$$\mathcal{F}_f^{A,B} := \{g \in \mathcal{F}^{A,B} / f^g \text{ is proper and } \inf \gamma_{g,f} = 0\}. \quad (8)$$

Remark: Observe that $\gamma_{g,f}$ might not be in $\mathcal{F}^{A,B}$, since $\gamma_{g,f}$ can take the value $+\infty$ for somewhere in $A \times B$.

Lemma 3.3 $\mathcal{F}_f^{A,B}$ is non-empty for all non-empty $B \subset Y$.

Proof: Given a non-empty $B \subset Y$, define $g \in \mathcal{F}^{A,B}$ by:

$$g(x, y) = \|y\|.$$

It is easy to check that g belongs to $\mathcal{F}_f^{A,B}$ (this example also proves that there are functions in $\mathcal{F}_f^{A,B}$ which satisfy (D2)).

Now consider

$$(P) \min_x f(x) \quad (9)$$

with $f \in \mathcal{F}^A$. Taking $g \in \mathcal{F}_f^{A,B}$, define the dual problem related to g :

$$(D_g) \min_{y \in B} f^g(y). \quad (10)$$

Since

$$\inf_{(x,y) \in A \times B} \gamma_{g,f}(x, y) = \inf_{x \in A} f(x) + \inf_{y \in B} f^g(y) = 0,$$

then

$$\inf_{x \in A} f(x) = - \inf_{y \in B} f^g(y) = \sup_{y \in B} [-f^g(y)]. \quad (11)$$

This means that there is no duality gap between the primal problem (P) and its dual (D_g) for every $g \in \mathcal{F}_f^{A,B}$.

Theorem 3.1 Let $g \in \mathcal{F}_f^{A,B}$. Then \bar{y} is a solution of (D_g) and \bar{x} is a solution of (P) if and only if $\gamma_{g,f}(\bar{x}, \bar{y}) = 0$.

Proof: \bar{x} and \bar{y} are solutions of (P) and (D_g) respectively if and only if

$$f(\bar{x}) = \inf f = - \inf f^g = -f^g(\bar{y}) \iff f(\bar{x}) + f^g(\bar{y}) = \gamma_{g,f}(\bar{x}, \bar{y}) = 0. \square$$

Remark: The previous result suggest us that the function $\gamma_{g,f}$ can be seen as the GAP function of problem (P) and its dual (D_g).

The next theorem states that given non-empty sets $A \subset X$ and $B \subset Y$, the correspondence defined by

$$\begin{aligned} \mathbf{F} : \mathcal{F}^A &\rightrightarrows \mathcal{F}^{A,B} \\ f &\mapsto \mathbf{F}(f) = \mathcal{F}_f^{A,B}, \end{aligned}$$

is a closed correspondence (see [9]).

Theorem 3.2 Take $f \in \mathcal{F}^A$ ($A \subset X$ is non-empty) and a non-empty $B \subset Y$. If there exist $f_k : \text{dom}(f) \rightarrow \mathbb{R}$, $g_k : A \times B \rightarrow \mathbb{R}$, sequences of functions ($k \in \mathbb{N}$), such that:

- i) f_k converges uniformly to f in $\text{dom}(f)$.
- ii) $g_k \in \mathcal{F}_{f_k}^{A,B}$ satisfies (D2) for every $k \in \mathbb{N}$.
- iii) g_k converges uniformly to a function g in $A \times B$.

Then $g \in \mathcal{F}_f^{A,B}$ and it satisfies (D2).

Proof: Let us prove first that $g \in \mathcal{F}^{A,B}$. Since g_k converges uniformly to g , given $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that if $k \geq N$ then

$$|g_k(x, y) - g(x, y)| < \varepsilon, \quad \forall (x, y) \in A \times B.$$

$$\text{Hence } g_k(x, y) - \varepsilon < g(x, y) < g_k(x, y) + \varepsilon, \quad \forall (x, y) \in A \times B.$$

Taking $\inf_{x,y}$ (remember that $\inf g_k = 0$ for all $k \in \mathbb{N}$):

$$-\varepsilon < \inf_{x,y} g(x, y) < \varepsilon.$$

Then $|\inf g| < \varepsilon$. And since $\varepsilon > 0$ is arbitrary, one has that $\inf g = 0$. This proves that $g \in \mathcal{F}^{A,B}$.

Now we prove that g satisfies (D2). We need to prove that $g(x, \cdot) : B \rightarrow \mathbb{R}$ is convex and l.s.c. for all $x \in A$. Let $x_0 \in A$ be fixed arbitrarily.

- $g(x_0, \cdot)$ is convex: since for all $k \in \mathbb{N}$, $g_k(x_0, \cdot)$ is convex, one has that given $y_1, y_2 \in B$ and $t \in [0, 1]$:

$$g_k(x_0, ty_1 + (1-t)y_2) \leq tg_k(x_0, y_1) + (1-t)g_k(x_0, y_2).$$

Making $k \rightarrow +\infty$:

$$g(x_0, ty_1 + (1-t)y_2) \leq tg(x_0, y_1) + (1-t)g(x_0, y_2),$$

which proves that $g(x_0, \cdot)$ is convex.

- $g(x_0, \cdot)$ is l.s.c.: fix $y_0 \in B$ and take $\lambda < g(x_0, y_0)$. There exists $N \in \mathbb{N}$ such that

$$|g_N(x, y) - g(x, y)| < \varepsilon, \quad \forall (x, y) \in A \times B,$$

$$\text{where } \varepsilon = \frac{g(x_0, y_0) - \lambda}{2}.$$

$$\text{Hence } \lambda < \lambda + \varepsilon = g(x_0, y_0) - \varepsilon < g_N(x_0, y_0).$$

Since $g_N(x_0, \cdot)$ is l.s.c., then there exists $V(y_0) \subset B$, a neighborhood of y_0 , such that if $y \in V(y_0)$ then

$$\lambda + \varepsilon < g_N(x_0, y).$$

Reducing $g(x_0, y)$:

$$\lambda + \varepsilon - g(x_0, y) < g_N(x_0, y) - g(x_0, y) < \varepsilon.$$

Therefore, if $y \in V(y_0)$, then $\lambda < g(x_0, y)$. Thus $g(x_0, \cdot)$ is l.s.c. in $y_0 \in B$, and since y_0 was fixed arbitrarily then $g(x_0, \cdot)$ is a l.s.c. function.

We have proved that for a fixed $x_0 \in A$, $g(x_0, \cdot)$ is a convex l.s.c. function, and since x_0 was fixed arbitrarily we have proved in fact that $g \in \mathcal{F}^{A,B}$ satisfies (D2).

It remains to prove that $g \in \mathcal{F}_f^{A,B}$. For doing this, let us show that $(f_k^{g_k})_{k \in \mathbb{N}}$ converges uniformly to f^g (in B).

Let $\varepsilon > 0$ and $N \in \mathbb{N}$ be such that if $k \geq N$ then

$$|g_k(x, y) - g(x, y)| < \frac{\varepsilon}{4}, \quad \forall (x, y) \in A \times B$$

and

$$|f_k(x) - f(x)| < \frac{\varepsilon}{4}, \quad \forall x \in \text{dom}(f).$$

Fix $k \geq N$ and take $y \in B$ arbitrarily, then

$$f_k^{g_k}(y) - \frac{\varepsilon}{2} < g_k(x', y) - f_k(x'), \quad \text{for some } x' \in \text{dom}(f).$$

Hence

$$f_k^{g_k}(y) - \varepsilon < g_k(x', y) - f_k(x') - \frac{\varepsilon}{2} < g(x', y) - f(x') \leq f^g(y),$$

and so

$$f_k^{g_k}(y) - \varepsilon < f^g(y). \quad (12)$$

This proves that $f_k^{g_k}(y) - f^g(y) < \varepsilon$. On the other hand:

$$f^g(y) - \frac{\varepsilon}{2} < g(x'', y) - f(x''), \quad \text{for some } x'' \in \text{dom}(f),$$

whence

$$f^g(y) - \varepsilon < g(x'', y) - f(x'') - \frac{\varepsilon}{2} < g_k(x'', y) - f_k(x'') \leq f_k^{g_k}(y),$$

and so

$$f^g(y) - \varepsilon < f_k^{g_k}(y).$$

This shows that

$$-\varepsilon < f_k^{g_k}(y) - f^g(y). \quad (13)$$

Since $y \in B$ was fixed arbitrarily, thanks to (12) and (13) we have that

$$-\varepsilon < f_k^{g_k}(y) - f^g(y) < \varepsilon, \quad \text{for every } y \in B.$$

This proves that $(f_k^{g_k})_{k \in \mathbb{N}}$ converges uniformly to f^g (in B), and it is immediate to see that f^g is proper and

$$0 \leq f(x) + f^g(y) \leq f_k(x) + f_k^{g_k}(y) + \varepsilon, \quad \forall (x, y) \in \text{dom}(f) \times B,$$

where $\varepsilon > 0$ is arbitrary and k is large enough. Taking $\inf_{(x,y) \in A \times B}$ one has:

$$0 \leq \inf_{(x,y) \in A \times B} (f(x) + f^g(y)) \leq \varepsilon.$$

Therefore $\inf_{(x,y) \in A \times B} (f(x) + f^g(y)) = 0$ and $g \in \mathcal{F}_f^{n,m}$. \square

This theorem proves a more difficult situation, the case when $g_k \in \mathcal{F}_{f_k}^{A,B}$ satisfy (D2) for all $k \in \mathbb{N}$. For the general case, just omit the two \bullet items and change B for a non-empty set.

At this point a natural question arises, for given $f \in \mathcal{F}^A$ and $g \in \mathcal{F}_f^{A,B}$, would be there any kind of relation between the optimal points and the optimal values of f and f^{gg} ? The next lemma answers this.

Lemma 3.4 For a fixed non-empty $B \subset Y$ and every $g \in \mathcal{F}_f^{A,B}$, the following are satisfied:

- i) $\inf f = \inf f^{gg}$,
- ii) if x_0 is a global minimum of f , then x_0 is a global minimum of f^{gg} .

Proof: Remember that f^{gg} is defined by:

$$f^{gg}(x) = \sup_{y \in B} \{g(x, y) - f^g(y)\}.$$

- i) $\inf f^{gg} \leq \inf f$ is always true. On the other hand

$$f^g(y) + f^{gg}(x) \geq g(x, y) \geq 0, \quad \forall x \in A, y \in B,$$

which implies that

$$\inf_{x \in A} f^{gg}(x) \geq - \inf_{y \in B} f^g(y).$$

But, since $g \in \mathcal{F}_f^{A,B}$ one has that

$$\inf f = - \inf_{y \in B} f^g(y),$$

which means

$$\inf f \leq \inf f^{gg} \leq \inf f.$$

Therefore $\inf f = \inf f^{gg}$.

- ii) $f^{gg}(x_0) \leq f(x_0) = \inf f = \inf f^{gg} \leq f^{gg}(x_0)$, then $f^{gg}(x_0) = \inf f^{gg}$.

4 Lagrangians induced by $\mathcal{F}_f^{A,B}$

Take $f \in \mathcal{F}^A$, a non-empty $B \subset Y$, $g \in \mathcal{F}_f^{A,B}$ and consider

$$(P) : \inf_{x \in A} f(x).$$

Recall that

$$(D_g) : \min_{y \in B} f^g(y)$$

is the dual problem of (P) related to g . Define $L_1 : \mathbb{R}^n \times C \rightarrow \mathbb{R} \cup \{+\infty\}$, as follows:

$$L_1(x, y) := f(x) - g(x, y). \tag{14}$$

This function has some interesting properties:

Theorem 4.1

$$\sup_{y \in B} \inf_{x \in A} L_1(x, y) = \inf_{x \in A} \sup_{y \in B} L_1(x, y). \tag{15}$$

Proof: The inequality $\sup_{y \in B} \inf_{x \in A} L_1(x, y) \leq \inf_{x \in A} \sup_{y \in B} L_1(x, y)$ is always true. For the opposite:

$$L_1(x, y) = f(x) - g(x, y) \leq f(x), \quad \forall (x, y) \in A \times B,$$

then

$$\sup_{y \in B} L_1(x, y) \leq f(x), \quad \forall x \in A.$$

It follows that

$$\inf_{x \in A} \sup_{y \in B} L_1(x, y) \leq \inf_{x \in A} f(x).$$

But, since $g \in \mathcal{F}_f^{A,B}$, we have that

$$\begin{aligned} \inf_{x \in A} f(x) &= - \inf_{y \in B} f^g(y) = - \left(\inf_{y \in B} \left\{ \sup_{x \in A} [g(x, y) - f(x)] \right\} \right) \\ &\implies \inf_{x \in A} f(x) = \sup_{y \in B} \inf_{x \in A} L_1(x, y), \end{aligned}$$

which means,

$$\inf_{x \in A} \sup_{y \in B} L_1(x, y) \leq \sup_{y \in B} \inf_{x \in A} L_1(x, y).$$

Finally,

$$\sup_{y \in B} \inf_{x \in A} L_1(x, y) = \inf_{x \in A} \sup_{y \in B} L_1(x, y). \square$$

We are interested now in which properties are satisfied for every saddle-point of L_1 . Remember that $(x_0, y_0) \in A \times B$ is a saddle point of L_1 if and only if

$$L_1(x_0, y) \leq L_1(x_0, y_0) \leq L_1(x, y_0), \quad \forall (x, y) \in A \times B.$$

Proposition 4.1 *Let L_1 be as before, if there exists $(x_0, y_0) \in A \times B$ saddle point of L_1 , then:*

- i) $x_0 \in \text{dom}(f)$.
- ii) y_0 is an optimal solution of (D_g) .
- iii) $f^{gg}(x_0) = f(x_0)$.

Proof:

- i) This is immediate thanks to the definition of saddle point.
- ii) From the previous theorem and the definition of saddle point, we have that

$$L_1(x_0, y_0) = \sup_{y \in B} \inf_{x \in A} L_1(x, y) = \inf_{x \in A} \sup_{y \in B} L_1(x, y).$$

But

$$\sup_{y \in B} \inf_{x \in A} L_1(x, y) = - \inf_{y \in B} f^g(y),$$

moreover

$$L_1(x_0, y_0) = \inf_{x \in A} L_1(x, y_0) = -f^g(y_0).$$

Thus,

$$f^g(y_0) = \inf_{y \in B} f^g(y).$$

- iii) $f^{gg}(x_0) = \sup_{y \in B} [g(x_0, y) - f^g(y)] = \sup_{y \in B} \left[g(x_0, y) - \sup_{z \in A} [g(z, y) - f(z)] \right]$. Which means,

$$f^{gg}(x_0) = \sup_{y \in B} \inf_{z \in A} [g(x_0, y) - g(z, y) + f(z)] = \sup_{y \in B} \inf_{z \in A} [g(x_0, y) + L_1(z, y)].$$

This implies

$$f^{gg}(x_0) \geq \inf_{z \in A} [g(x_0, y_0) + L_1(z, y_0)] = g(x_0, y_0) + \inf_{z \in A} L_1(z, y_0),$$

but since (x_0, y_0) is a saddle point of L_1 , then $\inf_{z \in A} L_1(z, y_0) = L_1(x_0, y_0)$. With this, we have that

$$f^{gg}(x_0) \geq g(x_0, y_0) + L_1(x_0, y_0) = f(x_0),$$

which means $f^{gg}(x_0) \geq f(x_0)$. $f^{gg}(x_0) \leq f(x_0)$ is always true (see [6] and references therein). \square

Proposition 4.2 *If x_0 is a solution of (P) and x_0^* is a solution of (D_g) , then (x_0, x_0^*) is a saddle point of L_1 .*

Proof: This is immediate from

$$0 \leq g(x_0, x_0^*) \leq f(x_0) + f^g(x_0^*) = 0 \implies g(x_0, x_0^*) = 0. \square$$

In Proposition 4.1 we would like to improve the fact that, in general, for every saddle point $(x_0, y_0) \in A \times B$ of L_1 we have that $f^{gg}(x_0) = f(x_0)$. For doing this, we impose an additional condition over g .

Proposition 4.3 *Let $g \in \mathcal{F}_f^{A,B}$ be such that $\inf_{y \in B} g(x, y) = 0$ for every $x \in A$. The following are equivalent:*

- i) (x_0, y_0) is a saddle-point of L .
- ii) x_0 is a solution of (P) and y_0 is a solution of (D_g) .

Proof: The implication ii) \Rightarrow i) is true thanks to the previous Proposition.

Consider now (x_0, y_0) a saddle-point of L_1 , then

$$L_1(x_0, y) \leq L_1(x_0, y_0), \quad \forall y \in B,$$

which is equivalent to

$$f(x_0) - g(x_0, y) \leq f(x_0) - g(x_0, y_0), \quad \forall y \in B$$

$$\Updownarrow$$

$$g(x_0, y_0) \leq g(x_0, y), \quad \forall y \in B.$$

Finally

$$g(x_0, y_0) = \inf_{y \in B} g(x_0, y) = 0.$$

On the other hand

$$L_1(x_0, y_0) \leq L_1(x, y_0), \quad \forall x \in A.$$

This implies that

$$f(x_0) \leq f(x) - g(x, y_0), \quad \forall x \in A$$

(remember that $g(x_0, y_0) = 0$). Taking $\inf_{x \in A}$ we have

$$f(x_0) \leq -f^g(y_0).$$

And thus $f(x_0) = -f^g(y_0)$, which means that x_0 is a solution of (P) and y_0 is a solution of (D_g) .

Remark: To prove that there exists a $g \in \mathcal{F}_f^{A,B}$ such that $\inf_{y \in B} g(x, y) = 0$ for every $x \in A$ just consider the trivial function $g \equiv 0$.

Examples

For these examples, consider $X = \mathbb{R}^n$, $h : X \rightarrow \mathbb{R}^m$,

$$A := \{x \in X : h(x) \leq 0\}$$

and $f : A \rightarrow \mathbb{R}$.

1. Classical Lagrangian

Let $Y = \mathbb{R}^m$ and h be such that $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex and l.s.c. for all $i = 1, \dots, m$. Consider

$$(CP) : \min_{x \in A} f(x),$$

where f is convex and l.s.c.

Remember that (see [1] and [3]) the following is the well known dual problem:

$$(D_L) : \min_{\lambda^* \geq 0} \sup_{x \in A} \{\langle \lambda^*, -h(x) \rangle - f(x)\},$$

$h(x) = (h_1(x), \dots, h_m(x))$. Moreover, x_0 is a solution of (CP) and λ_0^* is a solution of (D_L) if and only if (x_0, λ_0^*) is a saddle point of the Lagrangian function L , given by

$$L(x, \lambda^*) := f(x) + \langle \lambda^*, h(x) \rangle, \quad x \in A, \quad \lambda^* \in \mathbb{R}_+^m.$$

Taking $B := \mathbb{R}_+^m$, define $g : A \times B \rightarrow \mathbb{R}$ as follows:

$$g(x, \lambda^*) := \langle \lambda^*, -h(x) \rangle. \quad (16)$$

It is not difficult to show that $\mathcal{F}_f^{A,B}$ and, even more,

$$f^g(\lambda^*) = \sup_{x \in A} \{\langle \lambda^*, -h(x) \rangle - f(x)\}, \quad \lambda^* \in B.$$

Therefore, using G-coupling functions, we have recovered the classical lagrangian duality.

2. Non-linear lagrangian function

In [7] we find the following well studied case of a non-linear *lagrange-type* function:

$$L(x, \omega) = f(x) + \max\{\langle \omega_0, h(x) \rangle, \dots, \langle \omega_p, h(x) \rangle\},$$

where $x \in \mathbb{R}^n$ and $\omega \in (\mathbb{R}_+^m)^{1+p}$ ($p \in \mathbb{N}$).

If we consider $Y = (\mathbb{R}^m)^{1+p}$ and $B = (\mathbb{R}_+^m)^{1+p}$, define $g : A \times B \rightarrow \mathbb{R}$ as follows:

$$g(x, \omega) := \min(\langle -h(x), \omega_0 \rangle, \dots, \langle -h(x), \omega_p \rangle), \quad x \in A, \quad \omega \in B, \quad (17)$$

we will have that $g \in \mathcal{F}_f^{A,B}$ and the lagrangian function induced is the same Lagrange-type function given by [7].

Acknowledgements: I wish to thank to Dr. Wilfredo Sosa who was my supervisor during my Master thesis, most of these ideas are from that work. Special thanks to Dr. Regina Burachik, Dr. David Yost and Prof. Alex Rubinov for all their support and comments.

References

- [1] Avriel, Mordecai *Nonlinear Programming*. Prentice-Hall, INC., Englewood Cliffs, New Jersey (1976).
- [2] Burachik, Regina. *Generalized Proximal Point Methods for the Variational Inequality Problem*. Tese de Doutorado, IMPA. (1995).
- [3] Crouzeix, Jean Pierre; Ocaña, Eladio; Sosa, Wilfredo. *Análisis Convexo*. Monografía del IMCA (2003).
- [4] Morales-Silva, Daniel; Rubinov Alex; Sosa, Wilfredo. *G-coupling functions*. To appear in *Optimization*.
- [5] Rockafellar, R.T. *Convex Analysis*. Princeton University Press, Princeton, New Jersey (1970)
- [6] Rubinov, Alexander. *Abstract Convexity and Global Optimization*. Kluwer Academic Publishers, Dordrecht/Boston/London (2000).
- [7] Rubinov, Alexander and Yang Xiaoqi. *Lagrange-type Functions in Constrained non-convex Optimization*. Kluwer Academic Publishers, Dordrecht/Boston/London (2002).
- [8] Sosa Sandoval, Wilfredo. *Iterative Algorithms for the abstract Equilibrium Problem*. Tese de Doutorado, IMPA. (1999).
- [9] Zangwill, Willard *Nonlinear Programming. A unified Approach* Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1969).